DOI:10.11931/guihaia.gxzw201811035

两种梧桐叶绿体基因组密码子使用偏性分析

陆奇丰，骆文华，黄至欢*

（广西壮族自治区中国科学院广西植物研究所，广西 桂林 541006）

摘要：为分析美丽梧桐、云南梧桐叶绿体基因组密码子偏好性使用模式，本文通过筛选美丽梧桐、云南梧桐叶绿体基因组中的各 52 条蛋白编码序列，利用 CodonW、CUSP 和 SPSS 软件对其密码子使用模式及偏性进行分析。结果表明：美丽梧桐、云南梧桐的 GC 含量分别为 38.12%、38.05%，表明叶绿体基因组内富含 A/T 碱基。有效密码子数(ENC)范围为 36.91~56.46、36.55~58.04，多数密码子偏性较弱。相对同义密码子(RSCU)分析显示，RSCU>1 的密码子各有 29 个，其中 28 个以 A、U 结尾。中性绘图显示，GC3 与 GC12 的相关性不显著，回归曲线斜率分别为 0.195 和 0.304，说明密码子偏好性主要受到自然选择的影响。ENC-plot 分析中大部分基因分布于曲线的周围和下方，ENC 比值多分布于-0.04~0.10 之间，表明突变也影响密码子偏性的形成。此外，17、18 个密码子分别被鉴定为美丽梧桐、云南梧桐的最优密码子。美丽梧桐、云南梧桐叶绿体基因组的密码子使用偏性可能受选择和突变共同作用，且使用模式较为相似，但具有一定的差异，可能与适应环境的进化机制有关。

关键词：美丽梧桐，云南梧桐，叶绿体基因组，密码子偏性，最优密码子

中图分类号：Q943.2 **文献标识码：**A

Analysis on codon usage bias of chloroplast genome from two species in *Firmiana* Marsili

LU Qifeng, LUO Wenhua, HUANG Zhihuan*

(*Guangxi Institute of Botany, Guangxi Zhuang Autonomous Region and Chinese Academy of Sciences*, Guilin 541006, Guangxi, China)

Abstract: In order to analyze the codon usage patterns of the chloroplast genome of *Firmiana pulcherrima* and *Firmiana major*. Fifty-two genes selected from these two species of chloroplast genome, were analyzed using CodonW, CUSP and SPSS software. The result showed that the GC content in *Firmiana pulcherrima* was 38.12%, and *Firmiana major* was 38.05%, suggesting that A/T base was rich in chloroplast genome. The effective number of codons in *F. pulcherrima* ranged from 36.91 to 56.46, and *F. major* ranged from 36.55 to 58.04, meaning the codon bias was weak in *F. pulcherrima* and *F. major*. *F. pulcherrima* and *F. major* both had 29 codons with relative synonymous codon usage (RSCU) greater than 1 and 28 codons ending with A and U base.

基金项目：国家自然科学基金（31700196）；广西科学院基本业务费（2017YJJ23008）；广西植物研究所基本业务费（GZY16003，GZY17008）[Supported by the National Natural Science Foundation of China（31700196）；Fundamental Research Foundation of Guangxi Academy of Sciences （2017YJJ23008）；Fundamental Research Foundation of Guangxi Institute of Botany （GZY16003 , GZY17008）]

作者简介：陆奇丰（1988-），男，广西容县人，硕士，助理研究员，主要从事植物生物技术工作，（E-mail）luiqfeng@163.com。

***通信作者：**黄至欢，博士，副研究员，从事植物繁殖生态学工作，（E-mail）zhihuan_huang@whu.edu.cn。

Neutrality plot showed that there was not significant correlation between GC_3 and GC_{12}, and the regression slope was 0.195 in *F. pulcherrima*, and 0.304 in *F. major*, suggesting that codon bias was mainly affected by natural selection. ENC-plot analysis showed that most of genes fell below and followed the standard curve, and ENC ratio of most genes ranged from -0.04 to 0.10, suggesting that the mutation also affected the formation of codon bias in *F. pulcherrima* and *F. major*. In addition, 17 codons were identified as optimal codons in *F. pulcherrima*, and 18 in *F. major*. We concluded that the codon usage bias of *F. pulcherrima* and *F. major* are formed under effect of mutation and selection. Although the codon usage pattern of *F. pulcherrima* and *F. major* are very similar, there are some differences, which may be related to the evolutionary mechanism of adaptive environment.

Keywords: *Firmiana pulcherrima*, *Firmiana major*, chloroplast genome, codon usage bias, optimal codons

自然界中，基因序列根据三联密码子法则翻译成氨基酸序列，并进一步组装修饰形成具有特定空间结构和功能活性的蛋白质。密码子由 3 个相连的核苷酸碱基所组成，因此，一共存在 64 个密码子，除去 3 个终止密码子 (UAA、UAG、UGA)，61 个密码子对应特定的氨基酸，而天然氨基酸有 20 种，除甲硫氨酸(AUG)、色氨酸(UGG)对应单一密码子编码外，其余 18 种氨基酸均对应 2-6 个编码密码子，编码相同氨基酸的不同密码子又称为同义密码子，有研究发现，同义密码子使用频率在基因内并不是随机的，而是优先使用其中一些密码子，这一现象被称为密码子使用偏性(Gustafsson et al., 2004；Li et al., 2017)。影响密码子偏性的因素较多，如自然选择下 DNA 的变异(Ashesh, 2002；Suzuki, 2011)、tRNA 丰度(Buchan et al., 2006)、基因序列长度(Ribeiro et al., 2012；Gabriel et al., 2001；Stoletzki et al., 2011)、蛋白质翻译效率(Haupt et al., 2009；Li et al., 2017)等。研究密码子偏性不仅有助于解释物种间进化规律，同时对基因表达、载体构建、未知功能基因的分析具有重要意义。

叶绿体是植物细胞重要的细胞器之一，属于半自主性细胞器，是进行光合作用和固定 CO_2 的主要场所，还参与植物生长环境胁迫(干旱、盐、光等)相关的重要代谢过程(Liu et al., 2018)。此外，叶绿体具有母系遗传、结构稳定和基因组较小等特性，从而使其广泛运用于植物多样性、系统发育、DNA 条码开发、植物适应性和遗传工程研究中(Sajjad et al., 2017；Liu et al., 2018)。有研究发现，不同物种叶绿体基因组密码子使用偏性存在一定差异(Liu & Xue, 2005；郑薇玮等，2018)。因此，为提高叶绿体遗传转化效率及目的基因表达水平，极有必要开展叶绿体全基因组密码子使用模式的研究。

美丽梧桐(*Firmiana pulcherrima*)、云南梧桐(*Firmiana major*)是梧桐属重要物种，具有较高的研究价值和经济价值，由于生态环境的破坏，美丽梧桐、云南梧桐野外资源日益减少，

已被列为濒危树种(中国植物志，1984；周亚东等，2013；余志祥等，2013)。目前，对美丽梧桐、云南梧桐的研究重要集中于资源调查(黎国运等，2011；Yang et al., 2018)、种苗繁育(黎国运等，2011；周亚东等，2013；余志祥等，2013)等方面，虽然两种梧桐叶绿体全基因组已完成测序 (Ya et al., 2018；Wang et al., 2018)，但关于两者的叶绿体基因组编码基因的密码子使用模式比较的研究尚未见报道。因此，本文在前人的研究基础上，对美丽梧桐和云南梧桐的叶绿体基因组中 52 个基因的密码子碱基组成特点及使用偏性进行比较分析，揭示其密码子偏性的影响因素，分析最优密码子，旨在为美丽梧桐、云南梧桐制定全面的保育策略及为梧桐科叶绿体基因组学研究提供参考。

1 材料与方法

1.1 叶绿体基因序列数据

从 NCBI 分别检索并下载美丽梧桐、云南梧桐叶绿体全基因组序列信息，登录号分别为 NC_036395.1(*F. pulcherrima* 美丽梧桐)、NC_037242.1(*F. major* 云南梧桐)。分别获得 82 和 85 条编码基因，对获得的编码基因按以下的条件进行筛选：基因长度大于 300 bp；并以 ATG 为起始密码子及以 TAA、TAG、TGA 为终止密码子，并剔除重复的编码基因，最终获得各 52 条编码基因进行后续分析。

1.2 试验方法

1.2.1 密码子相关参数计算

利用 CodonW(http://codonw.sourceforge.net/)和 CUSP(http://imed.med.ucm.es/EMBOSS/)软件分析 CDS 的有效密码子数(ENC)、密码子第三位碱基组成特性，以 C_{T3}、C_{C3}、C_{A3}、C_{G3} 表示，统计密码子第 1、2、3 位位置碱基 G+C 及 52 个基因总 GC 含量，分别用 C_{GC1}、C_{GC2}、C_{GC3}、GC 表示。

1.2.2 中性绘图分析

计算各 CDS 的 C_{GC1}、C_{GC2} 的平均值，记为 C_{GC12}，并以此为纵坐标，以 C_{GC3} 为横坐标绘制散点图，分析 C_{GC3} 与 C_{GC12} 的相关性，并推测影响密码子偏性的因素。若 C_{GC3} 与 C_{GC12} 呈显著相关，表明密码子对碱基使用无差异，突变是密码子偏性的主要影响因素。若 C_{GC3} 与 C_{GC12} 相关性不显著，且斜率接近 0，表明密码子第 1、2 位碱基与第 3 位碱基使用存在差异，密码子偏性与第 3 位碱基显著相关，选择压力对密码子偏性的影响较强(尚明照等，2011；刘慧等，2017；秦政等，2018)。

1.2.3 PR2-plot 绘图分析

PR2-plot 分析是考察突变和自然选择对密码子偏性影响的方法之一(尚明照等，2011；刘慧等，2017)。为避免密码子第 3 位碱基上的 A/T 和 GC 突变的不平衡，通过统计两种梧

桐叶绿体 52 条基因中编码丙氨酸、甘氨酸、亮氨酸(CTT、CTC、CTA、CTG)、脯氨酸、精氨酸(CGT、CGC、CGA、CGG)、丝氨酸(TCT、TCC、TCA、TCG)、苏氨酸、缬氨酸的密码子第 3 位碱基的 T_3、C_3、A_3、G_3 的组成情况，并以 $A_3/(A_3+T_3)$为纵坐标，$G_3/(G_3+C_3)$为横坐标绘制散点图。

1.2.4 ENC-plot 绘图分析

为进一步了解两种梧桐碱基组成对密码子偏性的影响，以两种梧桐叶绿体中各基因的 ENC、GC3 值为纵、横坐标做散点图，并与 ENC 期望值($ENC = 2+GC3+29/([GC_3^2+ (1-GC_3)^2])$)进行比较，期望值代表密码子偏性仅受碱基组成的影响。若基因分布于期望值曲线的上方，其密码子偏性仅受突变影响；若基因分布于曲线下方较远位置，表明选择是影响密码子偏性的主要因素(Wright, 1990；牛元等，2018)。

1.2.5 相对同义密码子使用度分析

为考察同义密码子偏好性情况，以相对同义密码子使用度(RSCU)表示特定密码子在编码对应氨基酸的同义密码子间的相对使用概率(杨国锋等，2015)，当 RSCU>1 时，表示该密码子使用频率较高，偏好性较强；RSCU<1 时，表示该密码子使用频率较低，偏好性较弱；当 RSCU=1 时，该密码子无偏好性。

1.2.6 最优密码子分析

对 CDS 的 ENC 值进行排序，从两极各选出 5 个基因作为高、低表达组，分别计算对应的密码子 RSCU 值，筛选高低表达组内对应密码子△RSCU 之差大于 0.08 作为优越高表达密码子，并与整体密码子 RSCU 进行结合比较，将△RSCU>0.08，且整体 RSCU>1 的密码子定义为最优密码子(秦政等，2018；石元豹等，2018)。

2 结果与分析

2.1 密码子组成分析

通过 CodonW 和在线 CUSP 对美丽梧桐、云南梧桐叶绿体编码基因组的 52 条 CDS 进行分析，结果如表 1 所示：美丽梧桐、云南梧桐叶绿体编码基因密码子 C_{T3} 和 C_{A3} 的含量分别为 46.56%、46.67%与 43.04%、43.17%，表明两种梧桐叶绿体编码基因第三位碱基以 A/T 结尾为主。密码子上不同位置碱基 GC 含量有所差异，第一位碱基的 GC 含量最高，分别为 46.10%、46.17%；其次为第二位，分别为 38.14%、38.06%；最低为第三位碱基，分别为 30.11%、29.92%，且两种梧桐叶绿体基因组 GC 含量分别为 38.12%、38.05%，表明叶绿体编码基因富含 A/T 碱基。ENC 值可反映密码子偏性程度，其取值范围为 20~61，ENC 值越小，表明偏性越强，ENC 值越大，则密码子偏性越弱，一般以 ENC 的值为 35 时区分密码子偏性强弱(Jiang et al.，2008)。本研究中，两种梧桐叶绿体基因组的 ENC 值为 50 左右，大于 35，

表明其密码子偏性较弱。

利用在线 CUSP 对各基因密码子不同位置碱基的 GC 含量进行分析，结果见表 2，不同基因的 GC 含量存在一定的差异，C_{GC1}、C_{GC2} 与 C_{CG3} 差异较大，两种梧桐叶绿体编码基因 ENC 范围分别为 36.91~56.46、36.55~58.04，且大部分基因的 ENC 大于 44，表明叶绿体编码基因密码子偏性较弱。美丽梧桐、云南梧桐中 *atpE* 基因的 ENC 值最大，分别为 56.46 和 58.04，说明美丽梧桐中编码 *atpE* 基因密码子偏性强于云南梧桐，同时，两种梧桐叶绿体编码基因 GC 含量、ENC 值之间也存在差异，但差异不大。此外，本研究发现，美丽梧桐中不存在 *clpP* 基因及云南梧桐中的 *ndhK* 基因为假基因，表明两种梧桐叶绿体基因组具有较高的保守性，但种间存在一定的差异。

C_{GC1}、C_{GC2}、C_{CG3}、GC、ENC 及序列长度(sequence Length, SL)相关性分析结果见表 3，C_{GC1} 与 C_{GC2} 极显著相关，C_{GC1}、C_{GC2} 均与 C_{CG3} 相关性不显著，表明两种梧桐叶绿体 CDS 密码子的第 1、第 2 位碱基组成相似，且与第 3 位碱基组成存在较大差异，GC 与 C_{GC1}、C_{GC2}、C_{CG3} 均呈极显著相关，SL 与 C_{GC2} 呈负相关。此外，美丽梧桐的 ENC 与 C_{CG3} 呈显著相关，与 C_{GC1}、C_{GC2}、GC、SL 均未达到显著水平，表明美丽梧桐叶绿体编码基因密码子第 3 位碱基组成对密码子偏性影响较大。云南梧桐中 ENC 与 C_{GG1}、C_{GC3} 显著相关，相关值分别为 0.29、0.32，与 C_{GC2}、GC、SL 相关性较低，表明 C_{CG3} 对云南梧桐叶绿体密码子偏性的影响强于 C_{GC1}。两种梧桐密码子参数的相关性分析结果存在差异，其中 C_{GC1} 与云南梧桐叶绿体基因组密码子偏性显著相关，而美丽梧桐叶绿体基因组密码子偏性与 C_{GC1} 相关但不显著，进一步说明两种植物叶绿体基因组密码子偏性的影响因素有所差异。

表 1 两种梧桐叶绿体基因组密码子第 3 位碱基组成及 GC 含量

Table 1 GC contents and the third base of codon in chloroplast genes of *Firmiana pulcherrima* and *F. major* (%)

物种 Species	C_{T3}	C_{C3}	C_{A3}	C_{G3}	C_{GC1}	C_{GC2}	C_{GC3}	GC	ENC
美丽梧桐 *F. pulcherrima*	46.56	17.40	43.07	18.09	46.10	38.14	30.11	38.12	50.00
云南梧桐 *F. major*	46.67	17.23	43.17	18.03	46.17	38.06	29.92	38.05	49.94

表 2 两种梧桐叶绿体编码基因密码子 GC 含量(%)

Table 2 Genes GC contents in *Firmiana pulcherrima* and *F. major* chloroplast

基因 Gene	美丽梧桐 *F. pulcherrima*					云南梧桐 *F. major*				
	C_{GC1}	C_{GC2}	C_{GC3}	GC	ENC	C_{GC1}	C_{GC2}	C_{GC3}	GC	ENC
accD	40.37	35.50	26.57	34.14	47.29	40.37	35.29	26.57	34.08	47.11
atpA	54.72	40.16	28.15	41.01	47.63	54.92	40.35	28.15	41.14	47.84
atpB	56.11	41.28	31.86	43.09	48.20	56.51	41.48	31.86	43.29	48.50

atpE	52.24	40.30	34.33	42.29	56.46	52.24	40.30	35.07	42.54	58.04
atpF	47.57	32.97	34.05	38.20	47.85	47.57	32.97	34.05	38.20	46.83
atpI	50.00	37.50	28.63	38.71	45.14	50.00	37.50	28.63	38.71	45.61
ccsA	36.11	38.58	27.47	34.05	44.28	36.11	38.27	27.16	33.85	44.89
cemA	41.74	28.70	30.87	33.77	43.32	40.87	29.13	30.87	33.62	43.59
clpP			—			57.07	38.38	31.31	42.26	53.89
matK	39.36	31.21	28.83	33.33	49.32	39.76	31.21	27.83	32.94	49.79
ndhA	41.21	38.46	23.35	34.34	43.59	41.48	38.19	23.63	34.43	44.45
ndhB	41.88	39.14	31.31	37.44	47.84	41.88	39.14	31.31	37.44	47.84
ndhC	48.76	34.71	25.62	36.36	51.23	47.93	34.71	25.62	36.09	51.23
ndhD	40.86	37.85	30.32	36.34	49.72	40.86	37.85	29.68	36.13	49.23
ndhE	38.24	35.29	24.51	32.68	47.06	38.24	36.27	24.51	33.01	47.01
ndhF	36.77	35.55	25.51	32.61	46.61	36.91	35.28	24.83	32.34	45.68
ndhG	43.50	34.46	25.42	34.46	45.52	43.50	34.46	25.42	34.46	44.97
ndhH	50.00	36.29	28.68	38.32	48.67	50.00	36.80	28.17	38.32	49.08
ndhI	42.86	37.50	27.98	36.11	49.10	43.45	37.50	27.98	36.31	49.70
ndhJ	52.20	37.74	32.08	40.67	55.77	52.20	37.74	32.08	40.67	55.77
ndhK	44.25	45.13	28.32	39.23	51.27			Pseudogene		
petA	52.96	35.83	33.02	40.60	54.37	52.96	35.83	33.02	40.60	54.87
petB	49.54	41.20	31.48	40.74	45.78	49.54	41.20	31.02	40.59	45.57
petD	51.55	40.37	22.98	38.30	38.90	51.55	39.13	22.36	37.68	37.55
psaA	52.46	43.68	34.75	43.63	51.08	52.46	43.68	34.35	43.50	50.66
psaB	48.98	43.13	32.52	41.54	49.66	48.98	43.13	32.52	41.54	49.46
psbA	49.44	43.79	33.90	42.37	41.16	49.44	43.79	33.90	42.37	40.64
psbB	54.22	45.97	30.84	43.68	47.51	54.22	46.17	30.45	43.61	46.86
psbC	53.39	46.20	33.12	44.30	47.53	53.38	46.41	33.12	44.30	47.14
psbD	52.82	43.22	32.49	42.84	44.73	52.82	43.22	31.92	42.66	44.89
rbcL	57.44	43.19	31.66	44.10	50.22	57.11	43.30	31.55	43.99	50.39
rpl14	56.91	38.21	26.83	40.65	48.98	56.10	38.21	26.83	40.38	47.88
rpl16	48.53	52.21	27.21	42.65	40.54	48.53	52.21	27.94	42.89	40.54
rpl2	51.27	49.09	32.73	44.36	53.68	51.64	49.09	32.00	44.24	53.17
rpl20	38.98	44.92	24.58	36.16	47.40	38.98	43.22	25.42	35.88	48.82
rpl22	41.67	41.07	25.00	35.91	48.62	41.67	41.07	25.60	36.11	49.25
rpoA	45.43	32.01	28.35	35.26	49.64	45.12	32.01	28.05	35.06	49.04
rpoB	50.33	37.72	30.44	39.50	50.32	50.23	37.72	30.25	39.40	50.28
rpoC1	50.44	37.90	27.11	38.48	48.73	49.78	37.66	26.98	38.14	48.60
rpoC2	46.45	37.40	30.65	38.17	51.16	46.16	37.60	29.99	37.92	51.23
rps11	56.12	57.55	22.30	45.32	43.94	55.40	57.55	22.30	45.08	43.71
rps12	52.42	48.39	27.42	42.74	46.53	52.08	47.22	31.25	43.52	49.59
rps14	42.57	47.52	34.65	41.58	37.18	42.57	47.52	34.65	41.58	37.18
rps18	35.29	41.18	27.45	34.64	38.51	35.29	43.14	27.45	35.29	38.30
rps2	45.15	43.46	27.85	38.82	49.70	45.57	43.46	28.69	39.24	50.74
rps3	47.95	33.33	26.48	35.92	51.99	47.95	33.33	25.57	35.62	51.05
rps4	51.49	38.61	26.24	38.78	50.43	50.50	38.61	26.24	38.45	50.90

rps7	52.56	45.51	23.72	40.60	45.00	52.56	45.51	23.72	40.60	45.00
rps8	40.00	41.48	29.63	37.04	36.91	40.00	41.48	28.89	36.79	36.55
ycf1	36.53	29.30	25.52	30.45	48.07	36.62	29.13	25.09	30.28	47.94
ycf2	42.30	34.41	37.05	37.92	53.75	42.17	34.32	36.97	37.82	53.62
ycf3	47.93	38.46	33.14	39.84	55.71	47.93	38.46	32.54	39.64	53.19
ycf4	41.62	41.08	31.35	38.02	50.11	41.62	41.08	30.27	37.66	49.45

表 3 两种梧桐编码基因密码子相关参数的相关性分析

Table 3 Correlation coefficient between the parameters of genes

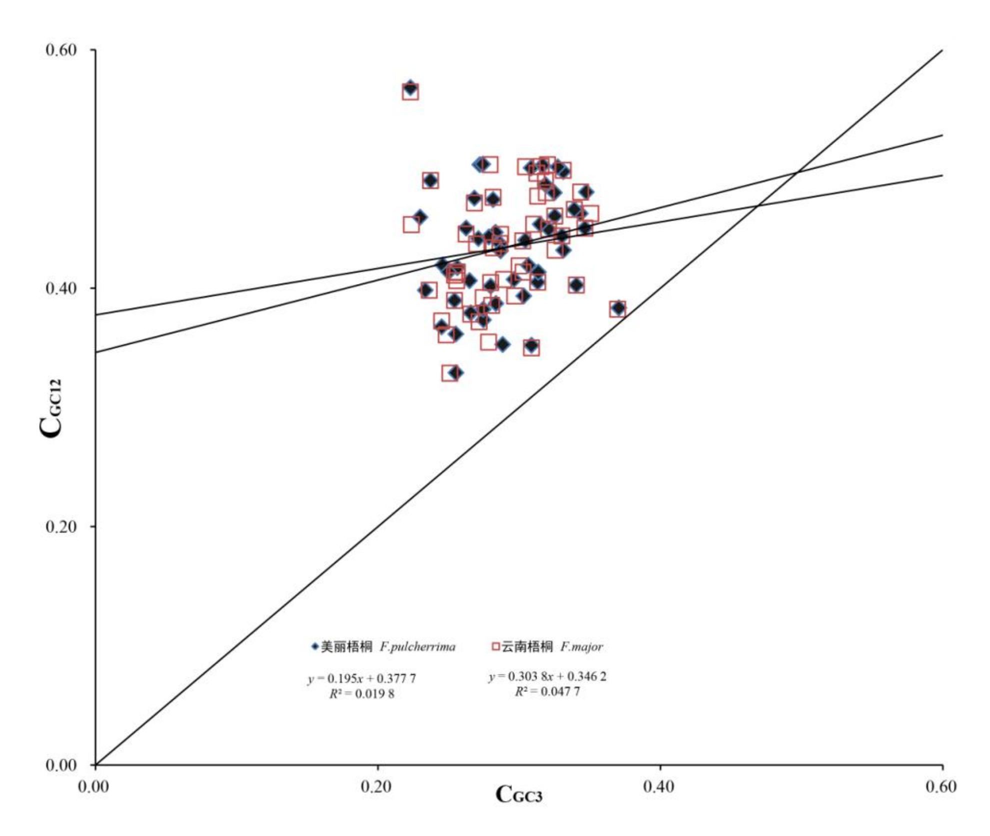
美丽梧桐 *F. pulcherrima*	云南梧桐 *F. major*					
	C_{GC1}	C_{GC2}	C_{GC3}	GC	ENC	SL
C_{GC1}	1.000	0.401**	0.273	0.841**	0.288*	-0.158
C_{GC2}	0.404**	1.000	0.083	0.747**	-0.257	-0.302*
C_{GC3}	0.222	0.005	1.000	0.509**	0.317*	0.214
GC	0.840**	0.743**	0.452**	1.000	0.133	-0.170
ENC	0.253	-0.261	0.313*	0.109	1.000	0.230
SL	-0.155	-0.309*	0.262	-0.161	0.220	1.000

注： *表示在 0.05 水平上显著相关； **表示在 0.01 水平上显著相关。

Note: * indicates significant correlation at 0.05 level, ** indicates significant correlation at 0.01 level.

2.2 中性绘图分析

如图 1 所示，美丽梧桐的叶绿体 CDS 的 C_{GC3} 与 C_{GC12} 含量分布分别为 0.223 0~0.370 5、0.329 2~0.568 4，云南梧桐为 0.223 0~0.369 7、0.328 8~0.564 8，两者的分布范围几乎重合，说明两种梧桐叶绿体基因组较为保守。两种梧桐的 C_{GC3} 与 C_{GC12} 的 Pearson 相关系数分别为 0.121 和 0.206，进一步说明 C_{GC1} 和 C_{GC2} 均与 C_{CG3} 相关性不显著，密码子的第 1、2 位碱基与第 3 位碱基组成存在显著差异，密码子偏性与第 3 位碱基显著相关。同时，回归曲线斜率分别为 0.195 和 0.304，表明选择压力对两种梧桐属叶绿体基因组密码子偏性的影响较强，此外，突变对密码子偏性也有一定的影响，且云南梧桐叶绿体密码子偏性受到突变的影响高于美丽梧桐。



注：C_{GC12} 表示密码子第 1、第 2 位碱基 GC 平均含量，C_{GC3} 表示第 3 位碱基 GC 含量。

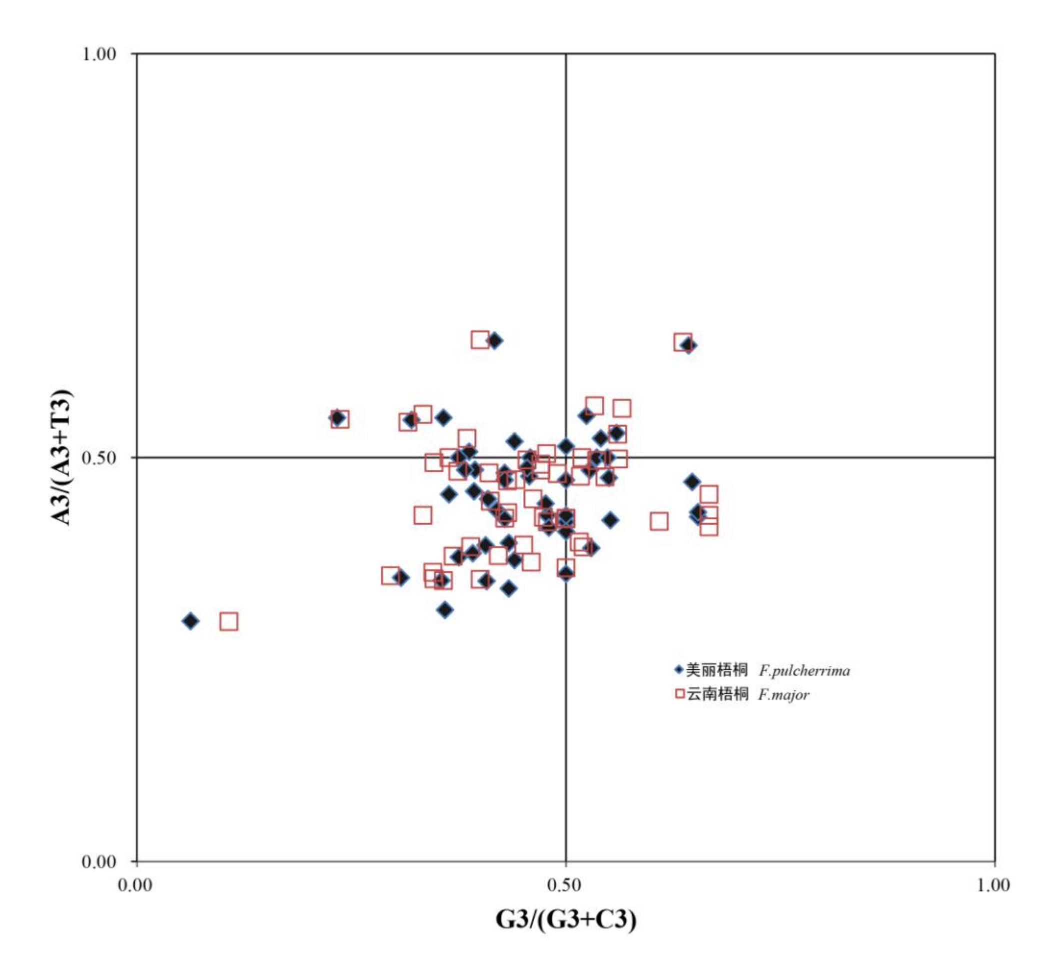
Note: C_{GC12} means average GC content of the first and the second position of these codons, C_{GC3} means the GC content of the third position of codons.

图 1 中性绘图分析

Fig.1 Analysis of neutrality plot C_{GC12} against C_{GC3}

2.3 PR2-plot 绘图分析

PR2-plot 绘图是通过对密码子第 3 位碱基组成进行分析，探讨突变和自然选择对密码子使用偏性的影响，当密码子第 3 位碱基表现为 A=T 和 C=G，表明 DNA 两条互补链上不存在选择效应上的偏性和突变，突变是影响密码子的偏性的唯一因素(Sueoka, 2001；Zhang et al., 2018)。本研究中，两种梧桐大部分基因分布在左下方区域(图 2)，T、C 碱基频率高于 A、G 碱基，说明密码子第 3 位碱基偏爱使用嘧啶碱基，表明两种梧桐叶绿体编码基因密码子使用偏性不仅受到突变的影响，还受到其他因素的影响，如自然选择压力等。



注：美丽梧桐横坐标平均位置为 0.453 2±0.073 9，纵坐标为 0.4503±0.103 7；云南梧桐横坐标平均位置为 0.454 4±0.073 6，纵坐标为 0.451 5±0.107 8。

Note: Open circles indicate the average position for each plot. *Firmiana pulcherrima*; average position is *x*=0.453 2±0.073 9, *y*=0.450 3±0.103 7; *F. major* average position is *x*=0.454 4±0.073 6, *y*=0.451 5±0.107 8.

图 2 PR2-plot 绘图分析

Fig.2 Analysis of PR2-plot [A3/(A3+T3)] against [G3/(G3+C3)]

2.4 ENC-plot 绘图分析

从图 3 中可以看出，两种梧桐的大部分基因集中分布于标准曲线的两侧，说明其 ENC 值与 ENC 期望值较为接近，而 *rps8*、*rps14*、*rps18*、*petD*、*rpl16*、*psbA* 基因偏离标准曲线较远，且 ENC 值较小，密码子偏性较强。根据公式(ENC $_{期望}$-ENC $_{实际}$)/ENC $_{期望}$统计基因 ENC 的分布频数(表 4)，结果显示：大部分基因的 ENC 频数集中于-0.04~0.10 之间，进一步表明大部分基因的 ENC 值与 ENC 期望值差异较小，表明这部分基因密码子偏性主要受到突变的影响。*rps8*、*rps14*、*rps18*、*petD*、*rpl16*、*psbA* 分布于 0.18~0.33 之间，与 0 点相差较远，表明这部分基因密码子偏性主要受到选择压力的影响。

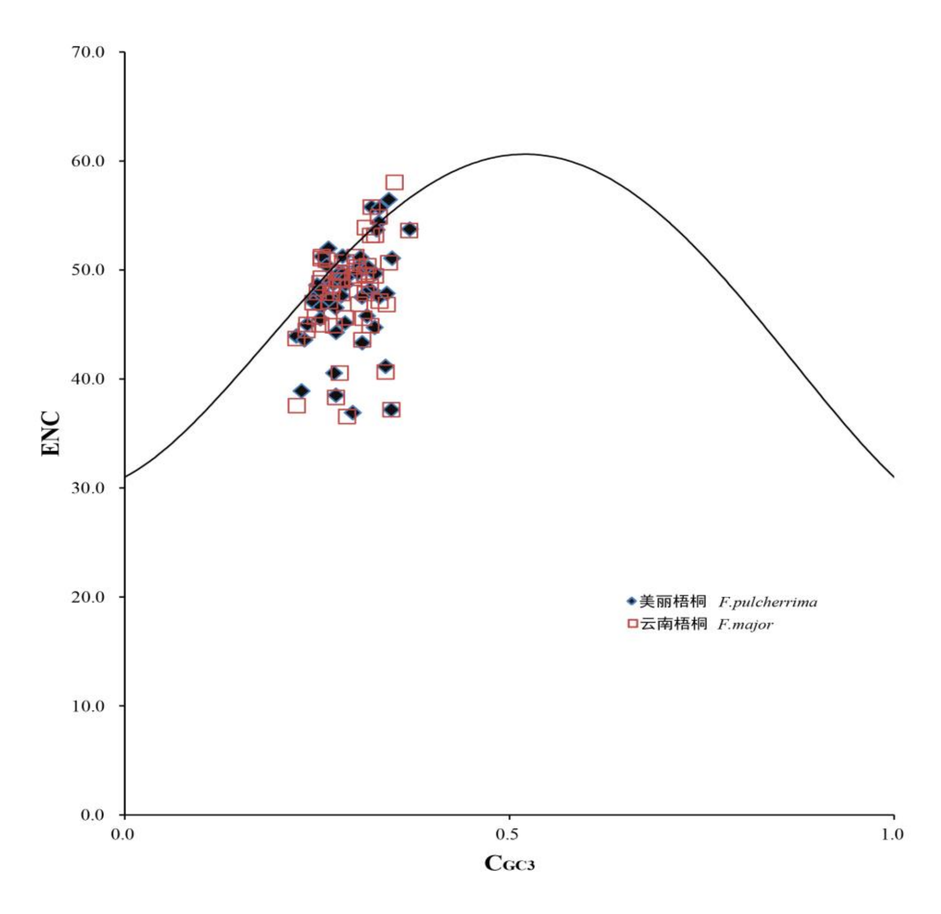


图 3 ENC-plot 绘图分析

Fig.3 Analysis of ENC-plot (ENC value against C_{GC3})

表 4 ENC 比值频率分布

Tab.4 Distribution of ENC frequency ratio

组限	组中值	美丽梧桐 *F. pulcherrima*		云南梧桐 *F. major*	
Class boundary	Class mid-value	组数 Number	组频 Frequency	频数 Number	组频 Frequency
-0.1~ 0	-0.05	9.00	0.17	9	0.17
0~0.1	0.05	30.00	0.58	29	0.56
0.1~0.2	0.15	9.00	0.17	10	0.19
0.2~0.3	0.25	3.00	0.06	3	0.06
0.3~0.4	0.35	1.00	0.02	1	0.02
总数 Total number		52.00	1.00	52	1.00

2.5 RSCU 分析

利用 CondonW 程序对两种梧桐叶绿体基因组 52 个编码基因进行 RSCU 分析，结果见表 5，从表中可以发现，两种梧桐的高频密码子(RSCU>1)各有 29 个，其中除 UUG(RSCU 大于 1)外，以 A/T 结尾的密码子 RSCU 值均在 1 以上，表明两种梧桐密码子偏好使用以 A/T 结尾的密码子，且两种梧桐植物对应密码子 RSCU 值极为接近，进一步说明其叶绿体基因

组密码子使用偏性较为接近，叶绿体基因组较为保守，种间差异较小。

表 5 两种梧桐密码子 RSCU 及最优密码子

Tab.5 Relative synonymous codon usage (RSCU) of genes and the optimal codons in *Firmiana pulcherrima* and *F. major* chloroplast

氨基酸 Amino Acid	密码子 Codon	美丽梧桐 *F. pulcherrima*			云南梧桐 *F. major*		
		叶绿体 *Cp* genome	高表达 High expression	低表达 Low expression	叶绿体 *Cp* genome	高表达 High expression	低表达 Low expression
Ala(A)	GCA*	1.09	1.14	0.90	1.08	1.14	0.99
	GCC	0.67	0.34	0.63	0.68	0.34	0.64
	GCG	0.47	0.80	0.70	0.48	0.80	0.60
	GCU	1.76	1.71	1.77	1.77	1.71	1.77
Cys(C)	UGC	0.50	0.40	0.59	0.52	0.40	0.63
	UGU*	1.50	1.60	1.41	1.48	1.60	1.37
Asp(D)	GAC	0.42	0.93	0.35	0.41	0.93	0.38
	GAU	1.58	1.07	1.65	1.59	1.07	1.62
Glu(E)	GAA*(*)	1.51	1.48	1.24	1.50	1.56	1.24
	GAG	0.49	0.52	0.76	0.50	0.44	0.76
Phe(F)	UUC	0.68	0.26	0.98	0.67	0.35	0.96
	UUU***	1.32	1.74	1.02	1.33	1.65	1.04
Gly(G)	GGA*	1.56	1.65	1.50	1.57	1.65	1.50
	GGC	0.43	0.35	0.47	0.43	0.35	0.51
	GGG	0.72	0.61	0.95	0.68	0.61	0.92
	GGU**	1.29	1.39	1.08	1.33	1.39	1.07
His(H)	CAC	0.51	0.40	0.50	0.50	0.40	0.51
	CAU*	1.49	1.60	1.50	1.50	1.60	1.49
Ile(I)	AUA	0.93	1.11	1.00	0.94	1.11	1.00
	AUC	0.57	0.32	0.74	0.57	0.32	0.71
	AUU(*)*	1.50	1.58	1.27	1.48	1.58	1.29
Lys(K)	AAA**	1.50	1.60	1.23	1.51	1.60	1.24
	AAG	0.50	0.40	0.77	0.49	0.40	0.76
Leu(L)	CUA	0.81	0.57	0.87	0.81	0.57	0.84
	CUC	0.37	0.19	0.58	0.37	0.19	0.59
	CUG	0.39	0.00	0.54	0.39	0.00	0.61
	CUU	1.23	1.52	1.44	1.25	1.52	1.50
	UUA***	1.97	2.19	1.10	1.96	2.19	0.99
	UUG	1.23	1.52	1.48	1.22	1.52	1.48
Met(M)	AUG	1.00	1.00	1.00	1.00	1.00	1.00
Asn(N)	AAC	0.47	0.72	0.48	0.46	0.69	0.48
	AAU	1.53	1.28	1.52	1.54	1.31	1.52

Pro(P)	CCA	1.13	0.78	1.22	1.12	0.67	1.24
	CCC	0.76	1.11	0.85	0.75	1.22	0.86
	CCG	0.56	0.78	0.64	0.57	0.78	0.56
	CCU	1.55	1.33	1.28	1.56	1.33	1.33
Gln(Q)	CAA***	1.55	2.00	1.37	1.54	2.00	1.42
	CAG	0.45	0.00	0.63	0.46	0.00	0.58
Arg(R)	AGA	1.77	1.94	1.76	1.80	1.92	1.89
	AGG	0.62	0.68	0.98	0.62	0.67	0.99
	CGA*(*)	1.39	1.61	1.47	1.40	1.67	1.36
	CGC	0.49	0.42	0.42	0.49	0.42	0.46
	CGG	0.43	0.25	0.65	0.40	0.25	0.56
	CGU**	1.29	1.10	0.72	1.29	1.08	0.73
Ser(S)	AGC	0.39	0.67	0.30	0.37	0.55	0.32
	AGU*	1.23	1.20	1.02	1.25	1.23	1.01
	UCA	1.18	0.93	1.14	1.20	0.95	1.18
	UCC	0.97	1.07	1.12	0.96	1.09	1.12
	UCG	0.54	0.80	0.81	0.55	0.82	0.76
	UCU	1.69	1.33	1.61	1.67	1.36	1.62
Thr(T)	ACA*	1.23	1.37	1.20	1.22	1.37	1.25
	ACC	0.80	1.03	0.77	0.78	0.91	0.69
	ACG	0.40	0.23	0.67	0.42	0.23	0.79
	ACU(*)	1.57	1.37	1.37	1.57	1.49	1.28
Val(V)	GUA***	1.46	2.32	1.24	1.47	2.32	1.23
	GUC	0.48	0.13	0.45	0.48	0.13	0.41
	GUG	0.59	0.13	0.82	0.60	0.13	0.82
	GUU	1.47	1.42	1.48	1.45	1.42	1.53
Trp(W)	UGG	1.00	1.00	1.00	1.00	1.00	1.00
Tyr(Y)	UAC	0.41	0.00	0.43	0.41	0.00	0.42
	UAU**	1.59	2.00	1.57	1.59	2.00	1.58
Codons No.	—	21138	635	3076	21124	635	3079

注：*表示ΔRSCU⩾0.08 且叶绿体基因组 RSCU>1；**表示ΔRSCU⩾0.3 且叶绿体基因组 RSCU>1；***表示ΔRSCU⩾0.5 且叶绿体基因组 RSCU>1，ΔRSCU 表示高、低表达基因库相应密码子之差。

Note: * means ΔRSCU⩾0.08 and chloroplast RSCU>1 ; **means ΔRSCU⩾0.3 and chloroplast RSCU>1; ***means ΔRSCU⩾0.5 and chloroplast RSCU>1, ΔRSCU means RSCU difference between highly expressed genes and low expressed genes

2.6 最优密码子分析

对两种梧桐植物叶绿体基因组基因的 ENC 值进行排序，选出两极的 10%基因作为高、低表达基因库，其中两种梧桐属叶绿体基因组高表达基因为 *rps8*、*rps14*、*petD*、*rps18*、*rpl16*，平均 ENC 值分别为 38.41(*F. pulcherrima*)、38.02(*F. major*)；*F. pulcherrima* 低表达基因为 *ycf2*、*petA*、*ycf3*、*ndhJ*、*atpE* ，而 *F. major* 为 *ycf2*、*clpP*、*petA*、*ycf3*、*ndhJ*、*atpE*，平均 ENC

值分别为 55.21(*F. pulcherrima*)、55.24(*F. major*)。分别计算高、低表达基因库的 RSCU 值，并计算对应密码子的△RSCU 值(高表达库 RSCU 值-低表达库 RSCU 值)，并与整体密码子的 RSCU 值结合比较，筛选出△RSCU 值≥0.08，且整体 RSCU>1 的密码子，定义为优越密码子。最终分别确定了 17、18 个表达优越密码子(见表 4)，分别为 GCA、UGU、GAA、UUU、GGA、GGU、CAU、AUU、AAA、UUA、CAA、CGA、CGU、AGU、ACA、ACU(为 *F. major* 优越密码子)、GUA、UAU，这些优越密码子均以 A/U 结尾。此外，我们观察到两种梧桐植物叶绿体基因组使用的优越密码子基本相同，但其△RSCU 值存在一定的差异，*F. pulcherrima* 中的 GAA、CGA 密码子△RSCU>0.08 而<0.30，而 *F. major* 中△RSCU>0.3，AUU 表现则相反，ACU 仅为 *F. major* 的优越密码子，在 *F. pulcherrima* 中△RSCU=0，推测可能与两种梧桐低表达库中所使用的编码基因差异有关。

3 讨论与结论

密码子是遗传信息的基本单位，是核酸与蛋白质的联系纽带，研究其使用情况对了解物种基因组进化具有重要意义(李平等，2011)。密码子使用偏性是物种不断进化适应的结果，其形成机制受多种因素的影响，其中突变作用和自然选择压力是其重要影响因素(王鹏良等，2018)。不同物种、基因家族中密码子使用偏性存在一定的差异，因此，密码子使用偏性是研究物种基因组进化关系及影响的重要指标之一。

碱基组成差异对密码子的使用有重要的影响。密码子在第 1、第 2 位碱基上的突变通常会导致编码氨基酸的改变，第 3 位碱基存在简并性，其突变对氨基酸改变通常无影响(尚明照等，2011)。因此，密码子第 3 位碱基的使用模式与密码子偏性密切相关。有研究表明，植物叶绿体基因组中密码子第三位碱基偏好使用 A/T(Zhou et al., 2008；秦政等，2018)。本研究中，两种梧桐植物 C_{T3} 和 C_{A3} 高于 C_{G3}、C_{C3}，且 C_{GC1}、C_{GC2} 均与 C_{CG3} 相关性不显著，C_{CG3} 比例最低，说明梧桐属植物叶绿体基因组密码子偏好使用 A/T 碱基结尾，与大花香水月季(*Rosa odorata* var. *gigantea*)、樟树(*Cinnamomum camphora*)、甜荞(*Fagopyrum esculentum*)(牛元等，2018；秦政等，2018；罗洪等，2015)等植物的研究结果一致。

自然选择和突变是影响密码子使用偏性的重要因素(Sharp et al.，2010；Rao et al.，2011)。本研究通过对两种梧桐叶绿体基因组 52 个编码基因 C_{GC3} 与 C_{GC12} 相关分析表明，密码子第 3 位碱基 GC 含量与第 1、第 2 位碱基相关性弱，且线性回归斜率较小，表明两种梧桐密码子的使用受到选择压力因素较强；此外，PR2-plot 分析发现，密码子第 3 位碱基上 A/T 碱基和 C/G 碱基使用频率存在差异，更偏好使用 T、C 碱基，进一步表明两种梧桐叶绿体基因组密码子使用不仅与碱基突变的有关，还更多受到选择的影响。ENC-plot 分析显示大部分基因的 ENC 值与期望 ENC 值较为接近，其分布频数集中于-0.04~0.10 之间，表明这部分基因

受密码子偏性与第 3 位碱基的 GC 含量相关，突变是主要影响因素；而 *rps8*、*rps14*、*rps18*、*petD*、*rpl16*、*psbA* 分布于标准曲线的下方较远位置，自然选择影响其密码子偏性的主要因素。综合中性绘图分析、ENC-plot、PR2-plot 等分析发现，突变和自然选择等多重因素共同影响着两种梧桐叶绿体密码子使用偏性。这与柿属植物、普通油茶、苦荞、甜荞等物种的叶绿体基因组密码子偏性研究结果相似(傅建敏等，2017；王鹏良等，2018；胡莎莎等，2016；罗洪等，2015)。

本研究发现，两种梧桐中 29 个高频密码子(RSCU>1)第 3 位碱基主要以 A/U 结尾，与前人在其他植物中的研究结果相似(尚明照等，2011；刘慧等，2017；秦政等，2018)。与高表达密码子进行分析，最终确定美丽梧桐的最优密码子为 17 个，云南梧桐 18 个，分别为 GCA、UGU、GAA、UUU、GGA、GGU、CAU、AUU、AAA、UUA、CAA、CGA、CGU、AGU、ACA、ACU(为 *F. major* 优越密码子)、GUA、UAU，这些最优密码子以 A/U 结尾，与大多植物叶绿体基因组最优密码子相似，表明梧桐与其他双子叶植物叶绿体基因组密码子第 3 位碱基偏爱使用 A/U 结尾(Kawabe & Miyashita, 2003)。有研究表明，物种间的亲缘关系越近，其密码子使用模式越相似(张太奎等，2017；张孟伟等，2018)，在本研究中，两种梧桐叶绿体基因组密码子使用偏性极为相似，但存在一定的差异。此外，研究还发现，美丽梧桐中叶绿体 *clpP* 编码基因丢失，云南梧桐的 *ndhK* 为假基因，推测两种梧桐适应环境改变的进化趋势有所不同，有待深入研究。

本研究分析比较美丽梧桐、云南梧桐叶绿体基因组编码基因的密码子偏性形成的影响因素和特点，以及两种梧桐叶绿体基因组的最优密码子，为外源基因密码子改造及梧桐属叶绿体基因组进化机制研究提供一定的理论基础。

参考文献

ASHESH N, 2002. Investigations on Evolutionary Changes in Base Distributions in Gene Sequences [J]. Internet Electr J Molec Design, 1: 545-558.

BUCHAN JR, AUCOTT LS, STANSFIELD I, 2006. tRNA properties help shape codon pair preferences in open reading frames [J]. Nucleic Acids Res, 34(3): 1015-1027.

CHINESE ACADEMY OF SCIENCE, 1984. Flora Reipublicae Popularis Sinicae [M], http://frps.iplant.cn/, 49(2): 136-137. [中国科学院，1984. 中国植物志[M]. http://frps.iplant.cn/，49(2): 136-137.]

FU JM, SUO YJ, LIU HM, et al., 2017. Analysis on codon usage in the chloroplast protein coding genes of *Diospyros* spp. [J]. Nonwood For Res, 35(2):38-44. [傅建敏，索玉静，刘慧敏，等，2017. 柿属植物叶绿体蛋白质编码基因密码子用法[J]. 经济林研究，35(2): 38-44.]

GABRIEL M, LAURENT D, 2001. Synonymous codon usage, accuracy of translation, and gene length in *Caenorhabditis elegans* [J]. J Mol Evol, 52: 275-280.

GUSTAFSSON C, GOVINDARAJAN S, MINSHULL J, 2004. Codon bias and heterologous

protein expression [J]. Trends Biotechnol, 22(7): 346-253.
HAUPT S, ZIEGLER A, COWAN G, et al., 2009. Studies of the role and function of barley stripe mosaic virus encoded proteins in replication and movement using GFP fusions [M]. Methods Mol Biol, 515: 287-297.
HU SS, LUO H, WU Q, et al., 2016. Analysis of codon usage bias of chloroplast genome of Tartary Buckwheat [J]. Mol Plant Breed, 14(2): 309-317. [胡莎莎，罗洪，吴琦，等，2016. 苦荞叶绿体基因组密码子偏爱性分析[J]. 分子植物育种，14(2): 309-317.]
JIANG Y, DENG F, WANG HL, et al., 2008. An extensive analysis on the global codon usage pattern of baculoviruses [J]. Arch Virol, 153(12): 2273-2282.
KAWABE and MIYASHITA, 2003. Patterns of codon usage bias in three dicot and four moncot plant species [J]. Genes Genet Syst, 78(5): 343-352.
LI GY, SHENG XH, XU PL, 2011. A preliminary report on wild resource survey of *Firmiana pulcherrima* [J]. Torp For, 39(1): 50-52. [黎国运，盛小彬，徐佩玲，2011. 美丽梧桐野生资源调查初报[J]. 热带林业，39(1): 50-52.]
LI GY, GUAN XJ, SHENG XH, et al., 2011. Research on shoot cutting propagation of *Firmiana pulcherrima Hsue* [J]. Torp For, 39(4): 36-38. [黎国运，关秀娟，盛小彬，等，2011. 美丽梧桐扦插育苗技术研究[J]. 热带林业，39(4): 36-38.]
LI J, LI HY, ZHI JK, et al., 2017. Codon usage of expansin genes in *Populus trichocarpa* [J]. Current Bioinform, 12: 452-461.
LI P, BAI YF, FENG RY, et al., 2011. Analysis of codon bias of NAD-ME Gene in *Amaranthus hypochondriacus* [J]. Chin J Appl Environ Biol, 1: 12-17. [李平，白云凤，冯瑞云，等，2011. 籽粒苋苹果酸酶(NAD-ME)基因密码子偏好性分析[J]. 应用与环境生物学报，1: 12-17.]
LIU H, WANG MX, YUE WJ, et al., 2017. Analysis of codon usage in the chloroplast genome of Broomcom millet (*Panicum miliaceum* L.) [J]. Plant Sci J, 35(3): 362-371. [刘慧，王梦醒，岳文杰，等，2017 糜子叶绿体基因组密码子使用偏性的分析[J]. 植物科学学报，35(3): 362-371.]
LIU QP, XUE QZ, 2005. Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species [J]. J Genet, 84(1): 55-62.
LIU X, ZHOU BY, YANG HY, et al., 2018. Sequencing and analysis of Chrysanthemum carinatum Schousb and Kalimeris indica. The complete chloroplast genomes reveal two inversions and rbcL as barcoding of the Vegetable [J]. Molecules, 23(6): 1358.
LUO H, HU SS, WU Q, et al., 2015. Analysis of Buckwheat chloroplast genome bias [J]. Genom Appl Biol, 34(11): 2457-2464. [罗洪，胡莎莎，吴琦，等，2015. 甜荞叶绿体基因密码子偏爱性分析[J]. 基因组学与应用生物学，34(11): 2457-2464.]
NIU Y, XU Q, WANG YD, et al., 2018. An analysis on codon usage bias of chloroplast genmoe of *Rosa odorata var.gigantea* [J]. J NW For Univ, 33(3): 123-130. [牛元，徐琼，王肴德，等，2018. 大花香水月季叶绿体基因组密码子使用偏性分析[J]. 西北林学院学报，33(3): 123-130].
QIN Z, ZHENG YJ, GUI LJ, et al., 2018. Codon usage bias analysis of the chloroplast genome of camphor tree (*Cinnamomum camphora*) [J]. Guihaia, 38(10): 1346-1355. [秦政，郑永杰，桂丽静，等. 樟树叶绿体基因组密码子偏好性分析[J]. 广西植物，38(10): 1346-1355.]
RAO Y, WU G, WANG Z, et al., 2011. Mutation bias is the driving force of codon usage in the gallus gallus genome [J]. DNA Res, 18(6): 499-512
RIBEIRO AS, HÄKKINEN A, LLOYD-PRICE J, 2012. Effects of gene length on the dynamics of

gene expression [J]. Comput Biol Chem, 41(4): 1-9.

SAJJAD ASAF, MUHAMMAD WAQAS, ABDUL L KHAN, et al., 2017. The complete chloroplast genome of wild rice (*Oryza minuta*) and Its comparison to related species [J]. Frontiers Plant Sci, 8: 304.

SHANG MZ, LIU F, HUA JP, et al., 2011. Analysis on codon usage of chloroplast genome of *Gossypium hirsutum* [J]. Sci Agric Sin, 44(2):245-253. [尚明照，刘方，华金平，等，2011. 陆地棉叶绿体基因组密码子使用偏性的分析[J]. 中国农业科学，44(2):245-253.]

SHARP PM EMERY LR, ZENG K, 2010. Forces that influence the evolution of codon bias [J]. Phil Trans R Soc B, 365(1544): 1203-1212.

SHI YB, WANG GB, YANG XM, et al., 2018. Analysis of codon usage of WRKY transcription factors in *Ginkgo biloba* [J]. Mol Plant Breeding, http:kns.cnki.net/kcms/detail/46.1068.S.20180809.1423.008.html. [石元豹，汪贵斌，杨晓明，等，2018. 银杏 WRKY 家族基因密码子使用偏向性分析[J]. 分子植物育种，http://kns.cnki.net/kcms/detail/46.1068.S.20180809.1423.008.html.]

STOLETZKI N, 2011. The surprising negative correlation of gene length and optimal codon use-disentangling translational selection from GC-biased gene conversion in yeast [J]. BMC Evol Biol, 11(1): 1179-1187.

SUEOKA N, 2001. Near homogeneity of PR2-bias fingerprintsin the human genome and their implications in phyloge-netic analyses [J]. J Mol Evol, 53: 469 –476.

SUZUKI Y, 2011. Statistical methods for detecting natural selection from genomic data [J]. Genes Genet Syst, 85(6): 359-376.

WANG JH, CAI YC, ZHAO KK, et al., 2018. Characterization of the complete chloroplast genome sequence of *Firmiana pulcherrima* (Malvaceae) [J]. Conserv Genet Resour, 10(3): 445–448.

WANG PL, YANG LP, WU HY, et al., 2018. Codon preference of chloroplast genome in *Camellia oleifera* [J]. Guihaia, 38(2): 135-144. [王鹏良，杨利平，吴红英，等，2018. 普通油茶叶绿体基因组密码子偏好性分析[J]. 广西植物，38(2): 135-144.]

WRIGHT F, 1990. The effective number of codons used in a gene [J]. Gene, 81(1): 23.

YA JD, YU ZX, YANG YQ, et al., 2018. Complete chloroplast genome of Firmiana major (Malvaceae), a critically endangered species endemic to southwest China [J]. Conserv Genet Resour, 10(4): 713-715.

YANG GF, SU KL, ZHAO YR, et al., 2015. Analysis of condon usage in the chloroplast genome of *Medicago truncatula* [J]. Atca Pratac Sin, 24(12): 171-179. [杨国锋，苏昆龙，赵怡然，等，2015. 蒺藜苜蓿叶绿体密码子偏好性分析[J]. 草业学报，24(12): 171-179.]

YANG J, CHEN G, SUN WB, 2018. Conserving *Firmiana major*, a tree species endemic to China [J]. Oryx, 52(2): 211- 213.

YU ZX, YANG YQ, MO X, et al., 2013. A shudy of artificial breeding of an endangered panlt (*Firmiana major*) [J]. J Sichuan For Sci Technol, 34(2): 62-65. [余志祥，杨永琼，莫旭，等，2013. 濒危植物云南梧桐(*Firmiana major*)人工繁育研究[J]. 四川林业科技，34(2): 62-65.]

ZHANG H, LI J, WANG RX, et al., 2018. Comparative analysis of expansin gene codonusage patterns among eight plant species[J]. J Biomol Struct Dyn, 7: 1-8.

ZHANG MW, ZHANG TK, LIU CY, et al., 2018. Biosysthsis pathway evolution and codon usage bias analysis of lignin in *Punica granatunm* [J]. Mol Plant Breed, 16(7): 2131-2138. [张孟

伟，张太奎，刘翠玉，等，2018. 石榴木质素合成途径进化与密码子使用偏好性分析[J]. 分子植物育种，16(7): 2131-2138.]
ZHANG TK, QI GH, YE HL, et al., 2017. Codon usage bias in pomegranate transcriptome [J]. Acta Hortic Sin, 44 (4): 675-690. [张太奎，起国海，叶红莲，等，2017. 石榴转录组密码子使用偏向性[J]. 园艺学报，44 (4): 675-690.]
ZHENG WW, CHEN JH, HAO ZD, et al., 2016. Analysis on codon usage bias of chloroplast genes from five conifers including *Cunninghamia lanceolata* [J]. Mol Plant Breed, 14(5): 1091-1097. [郑薇玮，陈金慧，郝兆东，等，2016. 杉木等 5 种针叶树叶绿体密码子偏好性分析[J]. 分子植物育种, 14(5): 1091-1097.]
ZHOU YD, ZHOU YB, YAN SL, et al., 2013. Effects of different substrates on seed germination and seeding growth of *Firmiana pulcherrima* Hsue [J]. J Trop Biol, 4(4): 322-326. [周亚东，周有标，颜速亮，等，2013. 不同基质对美丽梧桐种子育苗的影响[J]. 热带生物学报，4(4): 322-326.]